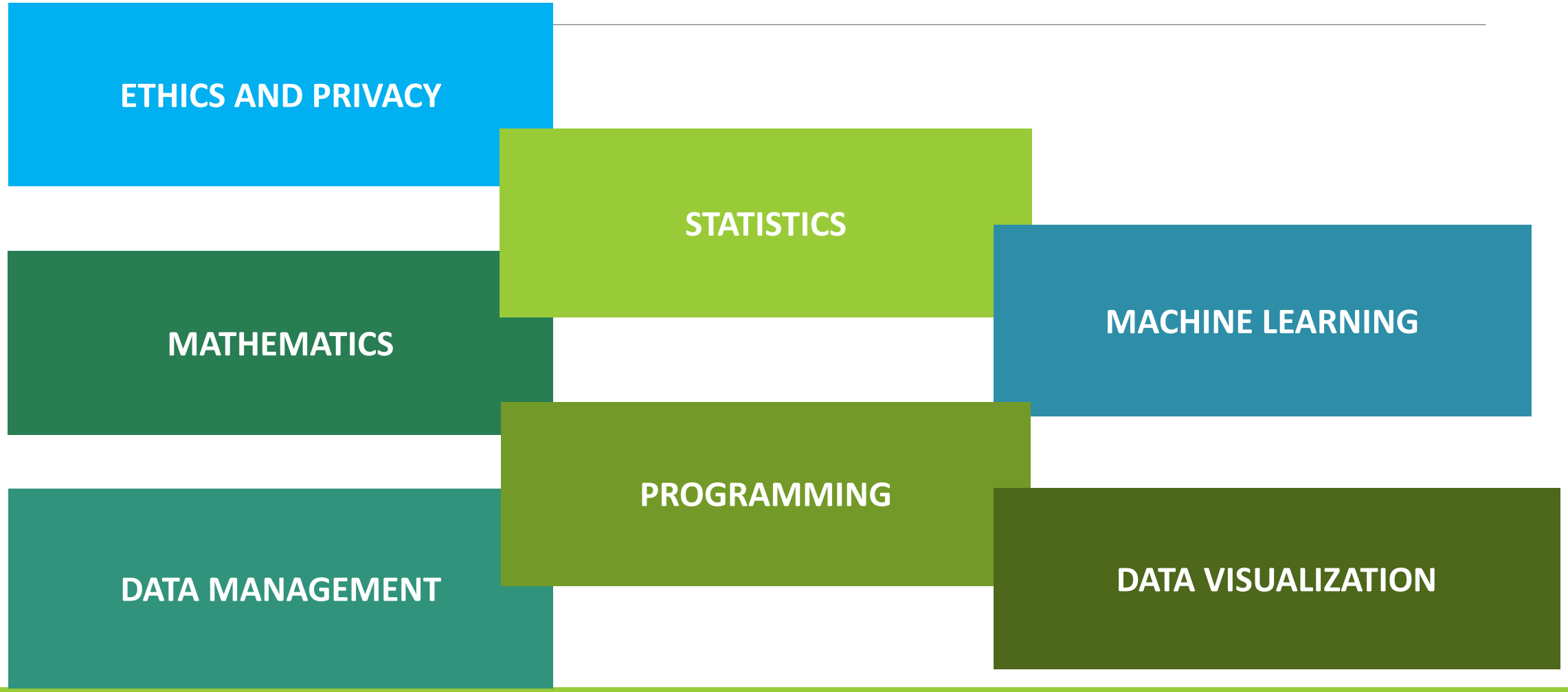
UN Big Data Competency Framework

Dominika Nowak Head of Experimental Research and Analyses Unit
Statistics Poland
do.nowak@stat.gov.pl

On behalf of the UN-CEBD Task Team on Training, Competencies and Capacity Development

UN Big Data Competency Framework - Core Competencies



UN Big Data Competency Framework - Soft Skills

AGILE PROJECT MANAGEMENT	ADAPTABILITY	BUSINESS ACUMEN
COMMUNICATION	CRITICAL THINKING	CURIOSITY
PRODUCT UNDERSTANDING	STORYTELLING	TEAM PLAYER

UN Big Data Competency Framework

Provides:

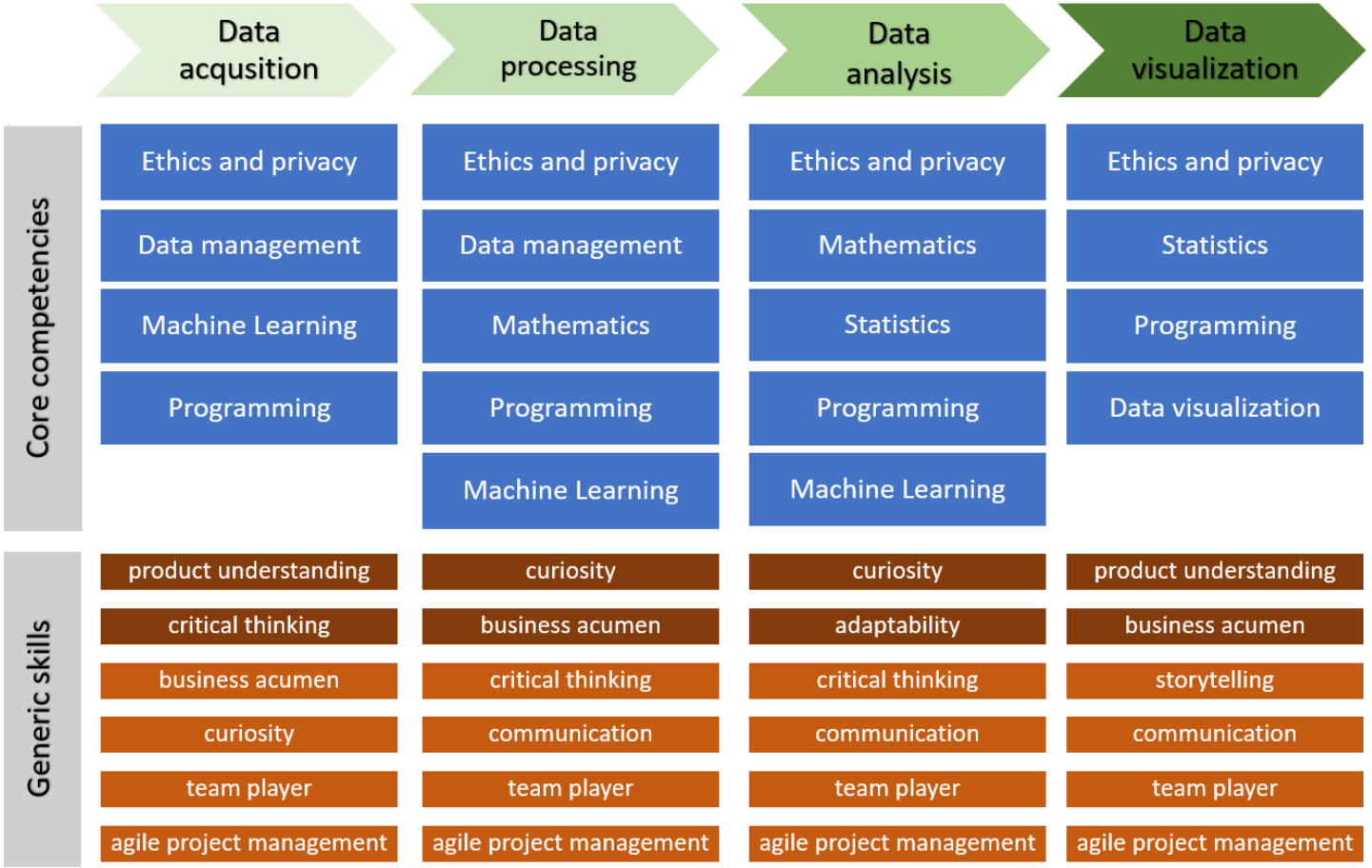
- General guidance on big data knowledge and skills
- Knowledge, skills and attitudes for acquiring and processing big data

Can be used to assess knowledge gaps, recruit and train staff at the NSO

Dimension 1			
Name of the area	Data management		
Dimension 2			
Competence title and description	To possess data management knowledge in a range of below-listed issues: 1) Database systems: database management systems, data models – definition and types, entity relationship model, models implementation (pre-relational, relational and object-oriented models) 2) Basics of cryptography: hash function, binary tree 3) Database: relational database, tabular data, data frames and series, shard, on-line analytical processing, data warehousing, data lakes, data vaults, logical multidimensional data model, extract, transform and load (ETL), NoSQL 4) Varied data formats: (Json, shp, XML, csv)		
Dimension 3	A - Foundation	B - Intermediate	C - Advanced
Proficiency levels	Demonstrate knowledge and understanding basic data management skills.	Demonstrate knowledge and understanding of, data base management tools and methods, and ability to apply some of them.	Thorough knowledge of proficiency in data base management and skillfulness in performing operations on varied data sets. Is able to advise others in finding data management solutions.
Dimension 4			
Knowledge examples	• Know the basic concept of SQL and NoSQL databases (such as table, column, field, field type, primary and foreign key) • Understand the consequences of using the hash function • Define functional dependencies occurring among the analyzed data		
Skills examples	• Able to create database structures in selected database management systems (e.g. MySQL, MongoDB, *more in annex*) • Able to present the logical structure of the database using tables and graphical relationships in selected programs • Apply ETL techniques - acquisition, processing (including pre-purification) and loading data from non-statistical sources		
Attitude examples	• Systematically supplement knowledge of new trends in the field of computer science on the subject of computer data storage • Identify data sources and assess their usefulness in complementing studies at hand • Carefully analyze the data and adjust them to the needs of database users		

Screen shot of Data Management section of UN Big Data Competency Framework

Big data competencies & statistical production process



UN Big Data Competency Framework

The Competency Framework should be used *as a guide*:

- Not every data specialist must possess all skills listed in the framework
- Different NSOs will run different projects that require different skills at different times
- Different types of data specialist (e.g. data analyst, data engineer, data scientist, etc.) require different skills and knowledge

How we can use the Competency Framework –
the case of Data Science Academy @Statistics Poland

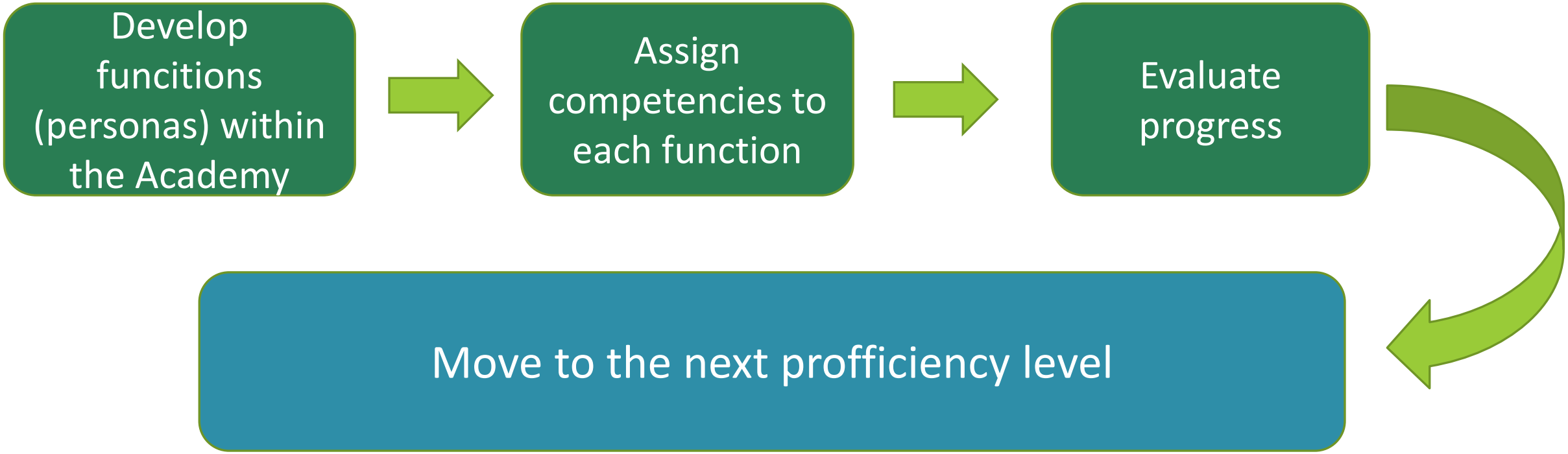
Internal developmnet program

Build a modern organization, able to anticipate and target fast-changing information needs by tapping into a multitude of data sources

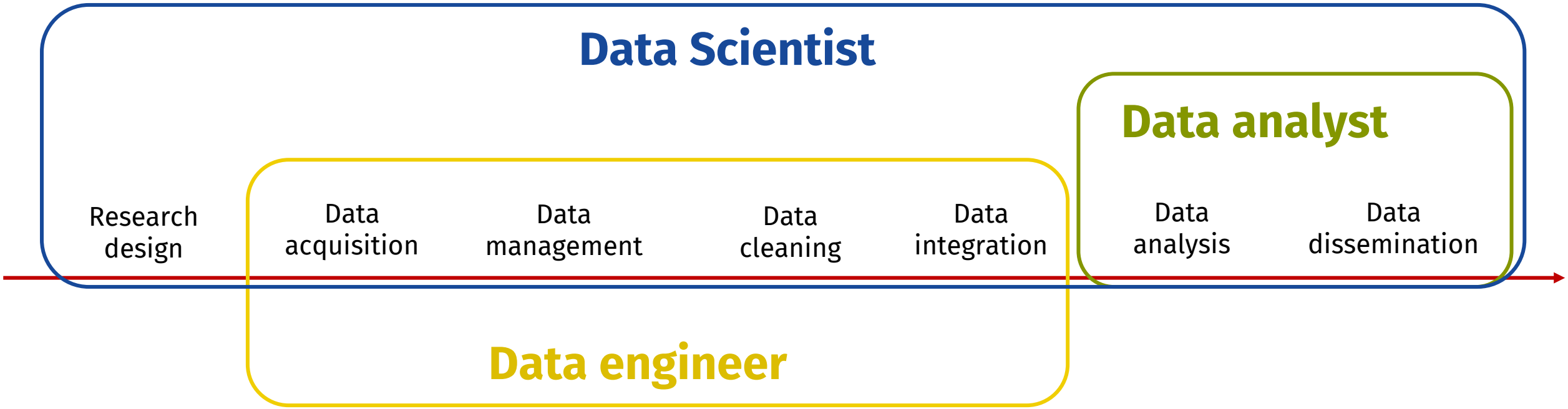
Increase the integration of administrative and big data sources in statistical production – experimental & official

How we can use the Competency Framework –
the case of Data Science Academy @Statistics Poland

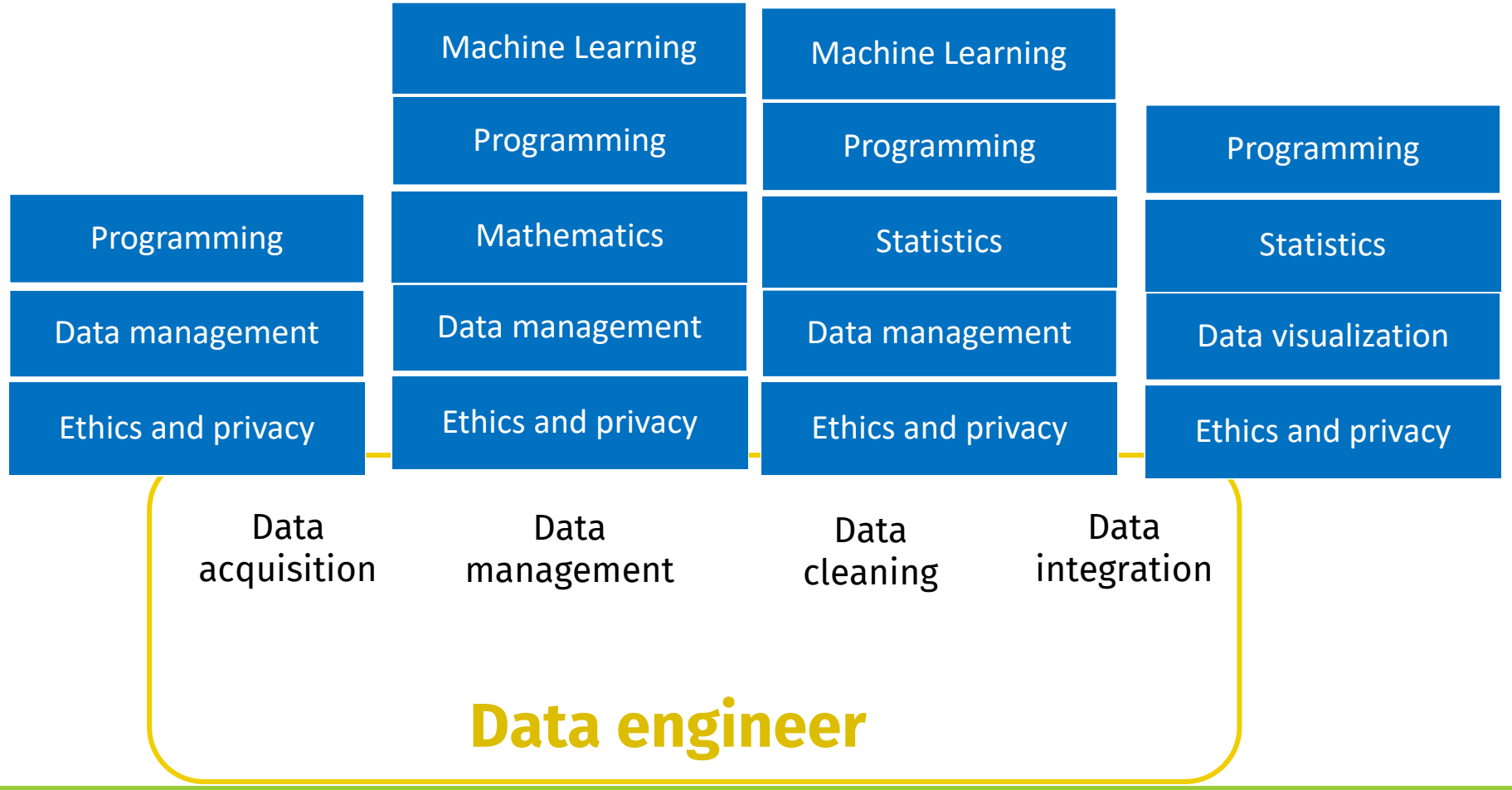
How we can use the Competency Framework –
the case of Data Science Academy @Statistics Poland



How we can use the Competency Framework –
the case of Data Science Academy @Statistics Poland



How we can use the Competency Framework –
the case of Data Science Academy @Statistics Poland



How we can use the Competency Framework –

the case of Data Science Academy @Statistics Poland

Progra

Data mar

Ethics an

Foundation

Intermediate

Advanced

Dimension 1	Name of the area	Data management		
Dimension 2				
	Competence title and description	To possess data management knowledge in a range of below-listed issues: 1) Database systems: database management systems, data models – definition and types, entity relationship model, models implementation (pre-relational, relational and object-oriented models) 2) Basics of cryptography: hash function, binary tree 3) Database: relational database, tabular data, data frames and series, shard, on-line analytical processing, data warehousing, data lakes, data vaults, logical multidimensional data model, extract, transform and load (ETL), NoSQL 4) Varied data formats: (json, shp, XML, csv)		
Dimension 3		A - Foundation	B - Intermediate	C - Advanced
	Proficiency levels	Demonstrate knowledge and understanding basic data management skills.	Demonstrate knowledge and understanding of, data base management tools and methods, and ability to apply some of them.	Thorough knowledge of proficiency in data base management and skillfulness in performing operations on varied data sets. Is able to advise others in finding data management solutions.
Dimension 4				
	Knowledge examples	• Know the basic concept of SQL and NoSQL databases (such as table, column, field, field type, primary and foreign key) • Understand the consequences of using the hash function • Define functional dependencies occurring among the analyzed data		
	Skills examples	• Able to create database structures in selected database management systems (e.g. MySQL, MongoDB, *more in annex*) • Able to present the logical structure of the database using tables and graphical relationships in selected programs • Apply ETL techniques - acquisition, processing (including pre-purification) and loading data from non-statistical sources		
	Attitude examples	• Systematically supplement knowledge of new trends in the field of computer science on the subject of computer data storage • Identify data sources and assess their usefulness in complementing studies at hand • Carefully analyze the data and adjust them to the needs of database users		

Data engineer

How we can use the Competency Framework –
the case of Data Science Academy @Statistics Poland

Data engineer	Data acquisition	Data Management	Data cleaning	Data integration
Foundation	Skill 1 👍	Skill 1 👍	Skill 1 👍	Skill 1 👍
	Skill 2 👍	Skill 2 👍	Skill 2 👍	Skill 2 👍
	Skill 3 👍		Skill 3 👍	
Intermediate	Skill 1 👍	Skill 1 👍	Skill 1 👍	Skill 1 👍
	Skill 2 👍	Skill 2 👍	Skill 2 👍	Skill 2 👍
	Skill 3 👍	Skill 3 👍	Skill 3 👍	Skill 3 👍
	Skill 4 👍			
Advanced	Skill 1	Skill 1	Skill 1	Skill 1
	Skill 2	Skill 2	Skill 2	Skill 2
	Skill 3		Skill 3	Skill 3

Entry assessment

Entry assessment

Entry assessment

Evaluation

Evaluation

Thank you!

Do.nowak@stat.gov.pl

UN Big Data Competency Framework